INGEGNERIA INFORMATICA (LM75)

(Lecce - Università degli Studi)

Insegnamento BIG DATA MANAGEMENT

GenCod A007906

Docente titolare ANTONELLA LONGO

Docenti responsabili dell'erogazione
ANTONELLA LONGO, MARCO SALVATORE ZAPPATORE

Insegnamento BIG DATA MANAGEMENT

Insegnamento in inglese BIG DATA MANAGEMENT

Settore disciplinare ING-INF/05

Corso di studi di riferimento
INGEGNERIA INFORMATICA

Tipo corso di studi Laurea Magistrale

Crediti 6.0

Ripartizione oraria Ore Attività frontale: 54.0

Per immatricolati nel 2024/2025

Erogato nel 2024/2025

Anno di corso 1

Lingua ITALIANO

Percorso PERCORSO COMUNE

Sede Lecce

Periodo Secondo Semestre

Tipo esame Orale

Valutazione Voto Finale

Orario dell'insegnamento
https://easyroom.unisalento.it/Orario

BREVE DESCRIZIONE DEL CORSO

The aim is of course is:
- to introduce datawarehouse concepts, methodologies and tools for the development of data analytics and OLAP systems
- to define the concept of big data and big data lifecycle
- to introduce methods and technologies to design and develop big data applications.

PREREQUISITI

Good knowledge of Object Oriented Languages (at least 1), techniques and tools. Elements of computer networks and Web technologies. EER models, Relational Databases and SQL are required.

OBIETTIVI FORMATIVI

Knowledge and understanding. Students must have a solid background related to the basics of big data management and information systems:

- They must have the basis to think analytically, creativelly and critically and being able to create abstraction and problem solving skills to cope with complex systems
- They must have a basic knowledge of design and implementation of big data management systems
- They must have the tools to design transactional and analytical databases applied to different contexts
- They must have the skills to argument data in different scenario, the tools for managing them, together with its impact.

Applying knowledge and understanding. After the course the student should be able to:

- Describe the model and frameworks of an Information System; illustrate the main components of an information system from the technical and application perspective.
- Distinguish conceptual, logical and physical models in big data management.
- Model Online Transaction processing systems from a big data perspective, distinguishing among conceptual models, relational models and physical models
- Model Online Analytical processing systems form a data perspective, distinguishing among conceptual, logical and physical models, being able to describe the relationships among them and the processes

Making judgements. Students are guided to critically approach the topics treated during the class, to compare different solutions to a problem, to identify and propose the most effective or efficient solution in an autonomous way.

Communication. Students must learn to communicate with heterogeneous audiences, explaining their position, in logical, coherent and effective way. During the course students will be provided with domain specific vocabulary and the proper scientific knowledge and methods to expose and argument in precise and formal way the main topics related to big data management and information systems

Learning skills. Students must acquire the critical ability to autonomously relate to the typical problems of data and information management and, in general, cultural issues related to information systems and their management. They should be able to develop an approach to independently structure knowledge and methods learnt with a view to possible continuation of studies at higher (doctoral) level or in the broader perspective of cultural and professional self-improvement of lifelong learning. Therefore, students should be able to switch their learning approach according to different learning sources and the objectives they must achieve in terms of results and audience

METODI DIDATTICI

The course aims to provide students with tools and knowledge for big data management in business organizations. The course consists of frontal lessons and classroom hands on exercises. The frontal lessons are aimed at improving students' knowledge and understanding through the presentation of theories, models and methods; students are invited to participate in the lesson with autonomy of judgement, by asking questions and presenting examples. The exercises are aimed at using tools which supports the models and approaches presented

MODALITA' D'ESAME	The exam consists of a written and oral part. The written part aims at assessing the modelling and design capabilites acquired by students. The oral part is based on the presentation of a lab activities and the interview will be related both on practical and descriptive aspects. The practical part aims at evaluating to what extent the student has: 1) the ability to design big data models according to the methodologies presented during the coursel, 2) reasoning about his/her choices and the capacity to integrate different concepts and tools. The descriptive part follows the practical part and is aimed to verify to what extent the student has gained knowledge and understanding of selected topics and he/she is able to communicate them.
ALTRE INFORMAZIONI UTILI	**Office Hours** By appointment; contact the instructor by email or at the end of class meetings.
PROGRAMMA ESTESO	▪ Data Warehouse Definition and lifecycle ▪ Modelling of analytical data processing systems: DFM, Snow Flakes schema ▪ Tools for data Visualization ▪ Introduction at big data and big data processing pipeline ▪ No SQL Databases and CAP Theorem ▪ Different kinds of no SQL Databases
TESTI DI RIFERIMENTO	Students who have not attended database course: Attend "Database Modeling and Design" by Prof. C. Batini on EDUOPEN: https://www.eduopen.org Main books: ▪R. Elmasri, S. Navathe, Fundamental of Database Systems, 7th Ed., Pearson ▪M. Golfarelli, S. Rizzi, Data Warehouse Design, Mc Graw Hill, 2021 ▪Thomas Erl and Wajid Khattak and and Paul Buhler. Big Data Fundamentals. Prentice Hall. 2016. ISBN 978-0-134-29107-9 Further references: ▪Joe Reis, Matt Housley, Fundamentals of Data Engineering, 2022, O'Relly ▪Balamurugan Balusamy, Nandhini Abirami R, Amir H. Gandomi, Big Data: Concepts, Technology, and Architecture, John Wiley & Sons Inc.; 1st Ed. ▪artin Kleppmann. Designing Data-Intensive Applications. O'Reilly. 2017. ISBN 978-1-449-37332-0 ▪Rajkumar Buyya and Rodrigo N. Calheiros and Amir Vahid Dastjerdi. Big Data: Principles and Paradigms. Morgan Kaufmann. 2016. 978-0-128-05394-2 ▪Ihab F. Ilyas and Xu Chu. 2019. Data Cleaning. Association for Computing Machinery, New York, NY, USA. Material provided during the class and references available on elearning.unisalento.it